

# SEQUENCE LISTING

<110> ITOH, Kyogo  
SHICHIJO, Shigeki  
IMAI, Yasuhisa

<120> TUMOR ANTIGEN PROTEINS, GENES THEREFOR, AND TUMOR  
ANTIGEN PEPTIDES

<130> 0020-4491P

<140> 09/202,047

<141> 1998-12-07

<160> 2

<170> PatentIn Ver. 2.1

<210> 1

<211> 2527

<212> DNA

<213> Homo sapiens

<220>

<221> 5'UTR

<222> (1)..(38)

<220>

<221> CDS

<222> (39)..(2438)

<220>

<221> 3'UTR

<222> (2439)..(2506)

<220>

<221> polyA\_site

<222> (2507)..(2527)

<400> 1

ggttcg'gcgcg cagccgggct cggagtggac gtgccact atg ggg tcg tcc aag aag 56  
Met Gly Ser Ser Lys Lys  
1 5

cat cgc gga gag aag gag gcg gcc ggg acg acg gcg gcg gcc ggc acc 104  
His Arg Gly Glu Lys Glu Ala Ala Gly Thr Thr Ala Ala Ala Gly Thr  
10 15 20

ggg ggt gcc acc gag cag ccg ccg cgg cac cgg gaa cac aaa aaa cac 152  
 Gly Gly Ala Thr Glu Gln Pro Pro Arg His Arg Glu His Lys Lys His  
 25 30 35

aag cac cgg agt ggc ggc agt ggc ggt agc ggt ggc gaa cga cgg aag 200  
 Lys His Arg Ser Gly Gly Ser Gly Gly Ser Gly Gly Glu Arg Arg Lys  
 40 45 50

cgg agc cgg gaa cgt ggg ggc gag cgc ggg agc ggg cgg cgc ggg gcc 248  
 Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly Ser Gly Arg Arg Gly Ala  
 55 60 65 70

gaa gct gag gcc cgg agc agc acg cac ggg cgg gag cgc agc cag gca 296  
 Glu Ala Glu Ala Arg Ser Ser Thr His Gly Arg Glu Arg Ser Gln Ala  
 75 80 85

gag ccc tcc gag cgg cgc gtg aag cgg gag aag cgc gat gac ggc tac 344  
 Glu Pro Ser Glu Arg Arg Val Lys Arg Glu Lys Arg Asp Asp Gly Tyr  
 90 95 100

gag gcc gct gcc agc tcc aaa act agc tca ggc gat gcc tcc tca ctc 392  
 Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser Gly Asp Ala Ser Ser Leu  
 105 110 115

agc atc gag gag act aac aaa ctc cgg gca aag ttg ggg ctg aaa ccc 440  
 Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala Lys Leu Gly Leu Lys Pro  
 120 125 130

ttg gag gtt aat gcc atc aag aag gag gcg ggc acc aag gag gag ccc 488  
 Leu Glu Val Asn Ala Ile Lys Lys Glu Ala Gly Thr Lys Glu Glu Pro  
 135 140 145 150

gtg aca gct gat gtc atc aac cct atg gcc ttg cga cag cga gag gag 536  
 Val Thr Ala Asp Val Ile Asn Pro Met Ala Leu Arg Gln Arg Glu Glu  
 155 160 165

ctg cgg gag aag ctg gcg gct gcc aag gag aag cgc ctg ctg aac caa 584  
 Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu Lys Arg Leu Leu Asn Gln  
 170 175 180

aag ctg ggg aag ata aag acc cta gga gag gat gac ccc tgg ctg gac 632  
 Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu Asp Asp Pro Trp Leu Asp  
 185 190 195

gac act gca gcc tgg atc gag agg agc cgg cag ctg cag aag gag aag 680  
 Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg Gln Leu Gln Lys Glu Lys  
 200 205 210

gac ctg gca gag aag agg gcc aag tta ctg gag gag atg gac caa gag 728  
 Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu Glu Glu Met Asp Gln Glu  
 215 220 225 230

ttt ggt gtc agc act ctg gtg gag gag gag ttc ggg cag agg cgg cag 776  
 Phe Gly Val Ser Thr Leu Val Glu Glu Glu Phe Gly Gln Arg Arg Gln  
 235 240 245

gac ctg tac agt gcc cgg gac ctg cag ggc ctc acc gtg gag cat gcc 824  
 Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly Leu Thr Val Glu His Ala  
 250 255 260

att gat tcc ttc cga gaa ggg gag aca atg att ctt acc ctc aag gac 872  
 Ile Asp Ser Phe Arg Glu Gly Glu Thr Met Ile Leu Thr Leu Lys Asp  
 265 270 275

aaa ggc gtg ctg cag gag gag gag gac gtg ctg gtg aac gtg aac ctg 920  
 Lys Gly Val Leu Gln Glu Glu Glu Asp Val Leu Val Asn Val Asn Leu  
 280 285 290

gtg gat aag gag cgg gca gag aaa aat gtg gag ctg cgg aag aag aag 968  
 Val Asp Lys Glu Arg Ala Glu Lys Asn Val Glu Leu Arg Lys Lys Lys  
 295 300 305 310

cct gac tac ctg ccc tat gcc gag gac gag agc gtg gac gac ctg gcg 1016  
 Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu Ser Val Asp Asp Leu Ala  
 315 320 325

cag caa aaa cct cgc tct atc ctg tcc aag tat gac gaa gag ctt gaa 1064  
 Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys Tyr Asp Glu Glu Leu Glu  
 330 335 340

ggg gag cgg cca cat tcc ttc cgc ttg gag cag ggc ggc acg gct gat 1112  
 Gly Glu Arg Pro His Ser Phe Arg Leu Glu Gln Gly Gly Thr Ala Asp  
 345 350 355

ggc ctg cgg gag cgg gag ctg gag gag atc cgg gcc aag ctg cgg ctg 1160  
 Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile Arg Ala Lys Leu Arg Leu  
 360 365 370

cag gct cag tcc ctg agc aca gtg ggg ccc cgg ctg gcc tcc gaa tac 1208  
 Gln Ala Gln Ser Leu Ser Thr Val Gly Pro Arg Leu Ala Ser Glu Tyr  
 375 380 385 390

ctc acg cct gag gag atg gtg acc ttt aaa aag acc aag cgg agg gtg 1256  
 Leu Thr Pro Glu Glu Met Val Thr Phe Lys Lys Thr Lys Arg Arg Val  
 395 400 405

aag aaa atc cgc aag aag gag aag gag gta gta gtg cgg gca gat gac 1304  
 Lys Lys Ile Arg Lys Lys Glu Lys Glu Val Val Val Arg Ala Asp Asp  
 410 415 420

ttg ctg cct ctc ggg gac cag act cag gat ggg gac ttt ggt tcc aga 1352  
 Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp Gly Asp Phe Gly Ser Arg  
 425 430 435

ctg cgg gga cgg ggt cgc cgc cga gtg tcc gaa gtg gag gag gag aag 1400  
 Leu Arg Gly Arg Gly Arg Arg Arg Val Ser Glu Val Glu Glu Glu Lys  
 440 445 450

gag cct gtg cct cag ccc ctg ccg tcg gac gac acc cga gtg gag aac 1448  
 Glu Pro Val Pro Gln Pro Leu Pro Ser Asp Asp Thr Arg Val Glu Asn  
 455 460 465 470

atg gac atc agt gat gag gag gaa ggt gga gct cca ccg ccg ggg tcc 1496  
 Met Asp Ile Ser Asp Glu Glu Glu Gly Gly Ala Pro Pro Pro Gly Ser  
 475 480 485

ccg cag gtg ctg gag gag gac gag gcg gag ctg gag ctg cag aag cag 1544  
 Pro Gln Val Leu Glu Glu Asp Glu Ala Glu Leu Glu Leu Gln Lys Gln  
 490 495 500

ctg gag aag gga cgc cgg ctg cga cag tta cag cag cta cag cag ctg 1592  
 Leu Glu Lys Gly Arg Arg Leu Arg Gln Leu Gln Gln Leu Gln Gln Leu  
 505 510 515

cga gac agt ggc gag aag gtg gtg gag att gtg aag aag ctg gag tct 1640  
 Arg Asp Ser Gly Glu Lys Val Val Glu Ile Val Lys Lys Leu Glu Ser  
 520 525 530

cgc cag cgg ggc tgg gag gag gat gag gat ccc gag cgg aag ggg gcc 1688  
 Arg Gln Arg Gly Trp Glu Glu Asp Glu Asp Pro Glu Arg Lys Gly Ala  
 535 540 545 550

atc gtg ttc aac gcc acg tcc gag ttc tgc cgc acc ttg ggg gag atc 1736  
 Ile Val Phe Asn Ala Thr Ser Glu Phe Cys Arg Thr Leu Gly Glu Ile  
 555 560 565

ccc acc tac ggg ctg gct ggc aat cgc gag gag cag gag gag ctc atg 1784  
 Pro Thr Tyr Gly Leu Ala Gly Asn Arg Glu Glu Gln Glu Glu Leu Met  
 570 575 580

gac ttt gaa cgg gat gag gag cgc tca gcc aac ggt ggc tcc gaa tct 1832  
 Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala Asn Gly Gly Ser Glu Ser  
 585 590 595

gac ggg gag gag aac atc ggc tgg agc acg gtg aac ctg gac gag gag 1880  
 Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr Val Asn Leu Asp Glu Glu  
 600 605 610

aag cag cag cag gat ttc tct gct tcc tcc acc acc atc ctg gac gag 1928  
 Lys Gln Gln Gln Asp Phe Ser Ala Ser Ser Thr Thr Ile Leu Asp Glu  
 615 620 625 630

gaa ccg atc gtg aat agg ggg ctg gca gct gcc ctg ctc ctg tgt cag 1976  
 Glu Pro Ile Val Asn Arg Gly Leu Ala Ala Ala Leu Leu Leu Cys Gln  
 635 640 645

aac aaa ggg ctg ctg gag acc aca gtg cag aag gtg gcc cgg gtg aag 2024  
 Asn Lys Gly Leu Leu Glu Thr Thr Val Gln Lys Val Ala Arg Val Lys  
 650 655 660

gcc ccc aac aag tcg ctg ccc tca gcc gtg tac tgc atc gag gat aag 2072  
 Ala Pro Asn Lys Ser Leu Pro Ser Ala Val Tyr Cys Ile Glu Asp Lys  
 665 670 675

atg gcc atc gat gac aag tac agc cgg agg gag gaa tac cga ggc ttc 2120  
 Met Ala Ile Asp Asp Lys Tyr Ser Arg Arg Glu Glu Tyr Arg Gly Phe  
 680 685 690

aca cag gac ttc aag gag aag gac ggc tac aaa ccc gac gtt aag atc 2168  
 Thr Gln Asp Phe Lys Glu Lys Asp Gly Tyr Lys Pro Asp Val Lys Ile  
 695 700 705 710

gaa tac gtg gat gag acg ggc cgg aaa ctc aca ccc aag gag gct ttc 2216  
 Glu Tyr Val Asp Glu Thr Gly Arg Lys Leu Thr Pro Lys Glu Ala Phe  
 715 720 725

cgg cag ctg tcg cac cgc ttc cat ggc aag ggc tca ggc aag atg aag 2264  
 Arg Gln Leu Ser His Arg Phe His Gly Lys Gly Ser Gly Lys Met Lys  
 730 735 740

aca gag cgg cgg atg aag aag ctg gac gag gag gcg ctc ctg aag aag 2312  
 Thr Glu Arg Arg Met Lys Lys Leu Asp Glu Glu Ala Leu Leu Lys Lys  
 745 750 755

atg agc tcc agc gac acg ccc ctg ggc acc gtg gcc ctg ctc cag gag 2360  
 Met Ser Ser Ser Asp Thr Pro Leu Gly Thr Val Ala Leu Leu Gln Glu  
 760 765 770

aag cag aag gct cag aag acc ccc tac atc gtg ctc agc ggc agc ggc 2408  
 Lys Gln Lys Ala Gln Lys Thr Pro Tyr Ile Val Leu Ser Gly Ser Gly  
 775 780 785 790

aag agc atg aac gcg aac acc atc acc aag tgacagcgcc ctcccgtagt 2458  
Lys Ser Met Asn Ala Asn Thr Ile Thr Lys  
795 800

cggccctgcc tcaaccttca tattaaataa agctccctcc ttatttttaa aaaaaaaaaa 2518

aaaaaaaaa 2527

<210> 2  
<211> 800  
<212> PRT  
<213> Homo sapiens

<400> 2  
Met Gly Ser Ser Lys Lys His Arg Gly Glu Lys Glu Ala Ala Gly Thr  
1 5 10 15

Thr Ala Ala Ala Gly Thr Gly Gly Ala Thr Glu Gln Pro Pro Arg His  
20 25 30

Arg Glu His Lys Lys His Lys His Arg Ser Gly Gly Ser Gly Gly Ser  
35 40 45

Gly Gly Glu Arg Arg Lys Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly  
50 55 60

Ser Gly Arg Arg Gly Ala Glu Ala Glu Ala Arg Ser Ser Thr His Gly  
65 70 75 80

Arg Glu Arg Ser Gln Ala Glu Pro Ser Glu Arg Arg Val Lys Arg Glu  
85 90 95

Lys Arg Asp Asp Gly Tyr Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser  
100 105 110

Gly Asp Ala Ser Ser Leu Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala  
115 120 125

Lys Leu Gly Leu Lys Pro Leu Glu Val Asn Ala Ile Lys Lys Glu Ala  
130 135 140

Gly Thr Lys Glu Glu Pro Val Thr Ala Asp Val Ile Asn Pro Met Ala  
145 150 155 160

Leu Arg Gln Arg Glu Glu Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu  
165 170 175

Lys Arg Leu Leu Asn Gln Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu  
180 185 190

Asp Asp Pro Trp Leu Asp Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg  
195 200 205

Gln Leu Gln Lys Glu Lys Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu  
210 215 220

Glu Glu Met Asp Gln Glu Phe Gly Val Ser Thr Leu Val Glu Glu Glu  
225 230 235 240

Phe Gly Gln Arg Arg Gln Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly  
245 250 255

Leu Thr Val Glu His Ala Ile Asp Ser Phe Arg Glu Gly Glu Thr Met  
260 265 270

Ile Leu Thr Leu Lys Asp Lys Gly Val Leu Gln Glu Glu Glu Asp Val  
275 280 285

Leu Val Asn Val Asn Leu Val Asp Lys Glu Arg Ala Glu Lys Asn Val  
290 295 300

Glu Leu Arg Lys Lys Lys Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu  
305 310 315 320

Ser Val Asp Asp Leu Ala Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys  
325 330 335

Tyr Asp Glu Glu Leu Glu Gly Glu Arg Pro His Ser Phe Arg Leu Glu  
340 345 350

Gln Gly Gly Thr Ala Asp Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile  
355 360 365

Arg Ala Lys Leu Arg Leu Gln Ala Gln Ser Leu Ser Thr Val Gly Pro  
370 375 380

Arg Leu Ala Ser Glu Tyr Leu Thr Pro Glu Glu Met Val Thr Phe Lys  
385 390 395 400

Lys Thr Lys Arg Arg Val Lys Lys Ile Arg Lys Lys Glu Lys Glu Val  
405 410 415

Val Val Arg Ala Asp Asp Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp  
420 425 430

Gly Asp Phe Gly Ser Arg Leu Arg Gly Arg Gly Arg Arg Arg Val Ser  
435 440 445

Glu Val Glu Glu Glu Lys Glu Pro Val Pro Gln Pro Leu Pro Ser Asp  
450 455 460

Asp Thr Arg Val Glu Asn Met Asp Ile Ser Asp Glu Glu Glu Gly Gly  
465 470 475 480

Ala Pro Pro Pro Gly Ser Pro Gln Val Leu Glu Glu Asp Glu Ala Glu  
485 490 495

Leu Glu Leu Gln Lys Gln Leu Glu Lys Gly Arg Arg Leu Arg Gln Leu  
500 505 510

Gln Gln Leu Gln Gln Leu Arg Asp Ser Gly Glu Lys Val Val Glu Ile  
515 520 525

Val Lys Lys Leu Glu Ser Arg Gln Arg Gly Trp Glu Glu Asp Glu Asp  
530 535 540

Pro Glu Arg Lys Gly Ala Ile Val Phe Asn Ala Thr Ser Glu Phe Cys  
545 550 555 560

Arg Thr Leu Gly Glu Ile Pro Thr Tyr Gly Leu Ala Gly Asn Arg Glu  
565 570 575

Glu Gln Glu Glu Leu Met Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala  
580 585 590

Asn Gly Gly Ser Glu Ser Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr  
595 600 605

Val Asn Leu Asp Glu Glu Lys Gln Gln Gln Asp Phe Ser Ala Ser Ser  
610 615 620

Thr Thr Ile Leu Asp Glu Glu Pro Ile Val Asn Arg Gly Leu Ala Ala  
625 630 635 640

Ala Leu Leu Leu Cys Gln Asn Lys Gly Leu Leu Glu Thr Thr Val Gln  
645 650 655

Lys Val Ala Arg Val Lys Ala Pro Asn Lys Ser Leu Pro Ser Ala Val  
660 665 670

Tyr Cys Ile Glu Asp Lys Met Ala Ile Asp Asp Lys Tyr Ser Arg Arg  
675 680 685



Glu Glu Tyr Arg Gly Phe Thr Gln Asp Phe Lys Glu Lys Asp Gly Tyr  
690 695 700

Lys Pro Asp Val Lys Ile Glu Tyr Val Asp Glu Thr Gly Arg Lys Leu  
705 710 715 720

Thr Pro Lys Glu Ala Phe Arg Gln Leu Ser His Arg Phe His Gly Lys  
725 730 735

Gly Ser Gly Lys Met Lys Thr Glu Arg Arg Met Lys Lys Leu Asp Glu  
740 745 750

Glu Ala Leu Leu Lys Lys Met Ser Ser Ser Asp Thr Pro Leu Gly Thr  
755 760 765

Val Ala Leu Leu Gln Glu Lys Gln Lys Ala Gln Lys Thr Pro Tyr Ile  
770 775 780

Val Leu Ser Gly Ser Gly Lys Ser Met Asn Ala Asn Thr Ile Thr Lys  
785 790 795 800

## SEQUENCE LISTING

SEQ ID NO: 1

SEQUENCE LENGTH: 800 amino acids

SEQUENCE TYPE: amino acid

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: human (*Homo sapiens*)

TISSUE TYPE: esophageal carcinoma tissue

10 FEATURE:

FEATURE KEY: peptide

LOCATION: 1..800

IDENTIFICATION METHOD: P

SEQUENCE DESCRIPTION:

15 Met Gly Ser Ser Lys Lys His Arg Gly Glu Lys Glu Ala Ala Gly Thr  
5 10 15  
Thr Ala Ala Ala Gly Thr Gly Gly Ala Thr Glu Gln Pro Pro Arg His  
20 25 30  
Arg Glu His Lys Lys His Lys His Arg Ser Gly Gly Ser Gly Gly Ser  
20 35 40 45  
Gly Gly Glu Arg Arg Lys Arg Ser Arg Glu Arg Gly Gly Glu Arg Gly  
50 55 60  
Ser Gly Arg Arg Gly Ala Glu Ala Glu Ala Arg Ser Ser Thr His Gly  
65 70 75 80  
25 Arg Glu Arg Ser Gln Ala Glu Pro Ser Glu Arg Arg Val Lys Arg Glu

85 90 95  
 Lys Arg Asp Asp Gly Tyr Glu Ala Ala Ala Ser Ser Lys Thr Ser Ser  
 100 105 110  
 Gly Asp Ala Ser Ser Leu Ser Ile Glu Glu Thr Asn Lys Leu Arg Ala  
 5 115 120 125  
 Lys Leu Gly Leu Lys Pro Leu Glu Val Asn Ala Ile Lys Lys Glu Ala  
 130 135 140  
 Gly Thr Lys Glu Glu Pro Val Thr Ala Asp Val Ile Asn Pro Met Ala  
 145 150 155 160  
 10 Leu Arg Gln Arg Glu Glu Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu  
 165 170 175  
 Lys Arg Leu Leu Asn Gln Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu  
 180 185 190  
 Asp Asp Pro Trp Leu Asp Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg  
 15 195 200 205  
 Gln Leu Gln Lys Glu Lys Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu  
 210 215 220  
 Glu Glu Met Asp Gln Glu Phe Gly Val Ser Thr Leu Val Glu Glu Glu  
 225 230 235 240  
 20 Phe Gly Gln Arg Arg Gln Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly  
 245 250 255  
 Leu Thr Val Glu His Ala Ile Asp Ser Phe Arg Glu Gly Glu Thr Met  
 260 265 270  
 Ile Leu Thr Leu Lys Asp Lys Gly Val Leu Gln Glu Glu Glu Asp Val  
 25 275 280 285  
 Leu Val Asn Val Asn Leu Val Asp Lys Glu Arg Ala Glu Lys Asn Val  
 290 295 300

Glu Leu Arg Lys Lys Lys Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu

305 310 315 320

Ser Val Asp Asp Leu Ala Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys

325 330 335

5 Tyr Asp Glu Glu Leu Glu Gly Glu Arg Pro His Ser Phe Arg Leu Glu

340 345 350

Gln Gly Gly Thr Ala Asp Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile

355 360 365

Arg Ala Lys Leu Arg Leu Gln Ala Gln Ser Leu Ser Thr Val Gly Pro

370 375 380

Arg Leu Ala Ser Glu Tyr Leu Thr Pro Glu Glu Met Val Thr Phe Lys

385 390 395 400

Lys Thr Lys Arg Arg Val Lys Lys Ile Arg Lys Lys Glu Lys Glu Val

405 410 415

15 Val Val Arg Ala Asp Asp Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp

420 425 430

Gly Asp Phe Gly Ser Arg Leu Arg Gly Arg Gly Arg Arg Val Ser

435 440 445

Glu Val Glu Glu Glu Lys Glu Pro Val Pro Gln Pro Leu Pro Ser Asp

20 450 455 460

Asp Thr Arg Val Glu Asn Met Asp Ile Ser Asp Glu Glu Glu Gly Gly

465 470 475 480

Ala Pro Pro Pro Gly Ser Pro Gln Val Leu Glu Glu Asp Glu Ala Glu

485 490 495

25 Leu Glu Leu Gln Lys Gln Leu Glu Lys Gly Arg Arg Leu Arg Gln Leu

500 505 510

Gln Gln Leu Gln Gln Leu Arg Asp Ser Gly Glu Lys Val Val Glu Ile

515 520 525  
 Val Lys Lys Leu Glu Ser Arg Gln Arg Gly Trp Glu Glu Asp Glu Asp  
 530 535 540  
 Pro Glu Arg Lys Gly Ala Ile Val Phe Asn Ala Thr Ser Glu Phe Cys  
 5 545 550 555 560  
 Arg Thr Leu Gly Glu Ile Pro Thr Tyr Gly Leu Ala Gly Asn Arg Glu  
 565 570 575  
 Glu Gln Glu Glu Leu Met Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala  
 580 585 590  
 10 Asn Gly Gly Ser Glu Ser Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr  
 595 600 605  
 Val Asn Leu Asp Glu Glu Lys Gln Gln Gln Asp Phe Ser Ala Ser Ser  
 610 615 620  
 Thr Thr Ile Leu Asp Glu Glu Pro Ile Val Asn Arg Gly Leu Ala Ala  
 15 625 630 635 640  
 Ala Leu Leu Leu Cys Gln Asn Lys Gly Leu Leu Glu Thr Thr Val Gln  
 645 650 655  
 Lys Val Ala Arg Val Lys Ala Pro Asn Lys Ser Leu Pro Ser Ala Val  
 660 665 670  
 20 Tyr Cys Ile Glu Asp Lys Met Ala Ile Asp Asp Lys Tyr Ser Arg Arg  
 675 680 685  
 Glu Glu Tyr Arg Gly Phe Thr Gln Asp Phe Lys Glu Lys Asp Gly Tyr  
 690 695 700  
 Lys Pro Asp Val Lys Ile Glu Tyr Val Asp Glu Thr Gly Arg Lys Leu  
 25 705 710 715 720  
 Thr Pro Lys Glu Ala Phe Arg Gln Leu Ser His Arg Phe His Gly Lys  
 725 730 735

Gly Ser Gly Lys Met Lys Thr Glu Arg Arg Met Lys Lys Leu Asp Glu

740

745

750

Glu Ala Leu Leu Lys Lys Met Ser Ser Ser Asp Thr Pro Leu Gly Thr

755

760

765

5 Val Ala Leu Leu Gln Glu Lys Gln Lys Ala Gln Lys Thr Pro Tyr Ile

770

775

780

Val Leu Ser Gly Ser Gly Lys Ser Met Asn Ala Asn Thr Ile Thr Lys

785

790

795

800

10 SEQ ID NO: 2

SEQUENCE LENGTH: 2527 base pairs

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

15 MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: No

ANTI-SENSE: No

ORIGINAL SOURCE:

ORGANISM: human (*Homo sapiens*)

20 TISSUE TYPE: esophageal carcinoma tissue

FEATURE:

FEATURE KEY: 5' UTR

LOCATION: 1..38

IDENTIFICATION METHOD: E

25

FEATURE KEY: CDS

LOCATION: 39..2438

IDENTIFICATION METHOD: E

5 FEATURE KEY: 3' UTR

LOCATION: 2439..2506

IDENTIFICATION METHOD: E

FEATURE KEY: poly A site

LOCATION: 2507..2527

IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION:

GGTTCGGCGG CAGCCGGGCT CGGAGTGGAC GTGCCACTAT GGGGTCGTCC AAGAAGCATC 60  
 GCGGAGAGAA GGAGGCGGCC GGGACGACGG CGGCGGCCGG CACCGGGGGT GCCACCGAGC 120  
 15 AGCCGCCGCG GCACCGGGAA CACAAAAAAC ACAAGCACCG GAGTGGCGGC AGTGGCGGTA 180  
 GCGGTGGCGA ACGACGGAAG CGGAGCCGGG AACGTGGGGG CGAGCGCGGG AGCGGGCGGC 240  
 GCGGGGCCGA AGCTGAGGCC CGGAGCAGCA CGCACGGGCG GGAGCGCAGC CAGGCAGAGC 300  
 CCTCCGAGCG GCGCGTGAAG CGGGAGAAGC GCGATGACGG CTACGAGGCC GCTGCCAGCT 360  
 CCAAACTAG CTCAGGCGAT GCCTCCTCAC TCAGCATCGA GGAGACTAAC AAACCTCCGGG 420  
 20 CAAAGTTGGG GCTGAAACCC TTGGAGGTTA ATGCCATCAA GAAGGAGGCG GGCACCAAGG 480  
 AGGAGCCCGT GACAGCTGAT GTCATCAACC CTATGGCCTT GCGACAGCGA GAGGAGCTGC 540  
 GGGAGAAGCT GGCGGCTGCC AAGGAGAAGC GCCTGCTGAA CCAAAGCTG GGGAAGATAA 600  
 AGACCCTAGG AGAGGATGAC CCCTGGCTGG ACGACACTGC AGCCTGGATC GAGAGGAGCC 660  
 GGCAGCTGCA GAAGGAGAAG GACCTGGCAG AGAAGAGGGC CAAGTTACTG GAGGAGATGG 720  
 25 ACCAAGAGTT TGGTGTGAGC ACTCTGGTGG AGGAGGAGTT CGGGCAGAGG CGGCAGGACC 780

TGTACAGTGC CCGGGACCTG CAGGGCCTCA CCGTGGAGCA TGCCATTGAT TCCTTCCGAG 840  
 AAGGGGAGAC AATGATTCTT ACCCTCAAGG ACAAAGGCGT GCTGCAGGAG GAGGAGGACG 900  
 TGCTGGTGAA CGTGAACCTG GTGGATAAGG AGCGGGCAGA GAAAAATGTG GAGCTGCGGA 960  
 AGAAGAAGCC TGA CTACCTG CCCTATGCCG AGGACGAGAG CGTGGACGAC CTGGCGCAGC 1020  
 5 AAAACCTCG CTCTATCCTG TCCAAGTATG ACGAAGAGCT TGAAGGGGAG CGGCCACATT 1080  
 CCTTCCGCTT GGAGCAGGGC GGCACGGCTG ATGGCCTGCG GGAGCGGGAG CTGGAGGAGA 1140  
 TCCGGGCCAA GCTGCGGCTG CAGGCTCAGT CCCTGAGCAC AGTGGGGCCC CGGCTGGCCT 1200  
 CCGAATACCT CACGCCTGAG GAGATGGTGA CCTTTAAAAA GACCAAGCAG AGGGTGAAGA 1260  
 AAATCCGCAA GAAGGAGAAG GAGGTAGTAG TGCGGGCAGA TGA CTGCTGCTG CCTCTCGGGG 1320  
 10 ACCAGACTCA GGATGGGGAC TTTGGTTCCA GACTGCGGGG ACGGGGTGCG CGCCGAGTGT 1380  
 CCGAAGTGGA GGAGGAGAAG GAGCCTGTGC CTCAGCCCCT GCGGTCCGAC GACACCCGAG 1440  
 TGGAGAACAT GGACATCAGT GATGAGGAGG AAGGTGGAGC TCCACCGCCG GGGTCCCCGC 1500  
 AGGTGCTGGA GGAGGACGAG GCGGAGCTGG AGCTGCAGAA GCAGCTGGAG AAGGGACGCC 1560  
 GGCTGCGACA GTTACAGCAG CTACAGCAGC TGCGAGACAG TGGCGAGAAG GTGGTGGAGA 1620  
 15 TTGTGAAGAA GCTGGAGTCT CGCCAGCGGG GCTGGGAGGA GGATGAGGAT CCCGAGCGGA 1680  
 AGGGGGCCAT CGTGTTCAAC GCCACGTCCG AGTTCTGCCG CACCTTGGGG GAGATCCCCA 1740  
 CCTACGGGCT GGCTGGCAAT CGCGAGGAGC AGGAGGAGCT CATGGACTTT GAACGGGATG 1800  
 AGGAGCGCTC AGCCAACGGT GGCTCCGAAT CTGACGGGGA GGAGAACATC GGCTGGAGCA 1860  
 CGGTGAACCT GGACGAGGAG AAGCAGCAGC AGGATTTCTC TGCTTCCTCC ACCACCATCC 1920  
 20 TGGACGAGGA ACCGATCGTG AATAGGGGGC TGGCAGCTGC CCTGCTCCTG TGTCAGAACA 1980  
 AAGGGCTGCT GGAGACCACA GTGCAGAAGG TGGCCCGGGT GAAGGCCCCC AACAAGTCGC 2040  
 TGCCCTCAGC CGTGTACTGC ATCGAGGATA AGATGGCCAT CGATGACAAG TACAGCCGGA 2100  
 GGGAGGAATA CCGAGGCTTC ACACAGGACT TCAAGGAGAA GGACGGCTAC AAACCCGACG 2160  
 TTAAGATCGA ATAGGTGGAT GAGACGGGCC GGAAACTCAC ACCCAAGGAG GCTTTCCGGC 2220  
 25 AGCTGTCGCA CGCTTCCAT GGCAAGGGCT CAGGCAAGAT GAAGACAGAG CGGCGGATGA 2280  
 AGAAGCTGGA CGAGGAGGCG CTCCTGAAGA AGATGAGCTC CAGCGACACG CCCCTGGGCA 2340  
 CCGTGGCCCT GCTCCAGGAG AAGCAGAAGG CTCAGAAGAC CCCCTACATC GTGCTCAGCG 2400



GCAGCGGCAA GAGCATGAAC GCGAACACCA TCACCAAGTG ACAGCGCCCT CCCGTAGTCG 2460  
GCCCTGCCTC AACCTTCATA TTAAATAAAG CTCCTCCTT ATTTTAAAA AAAAAAAAAA 2520  
AAAAAAA 2527

GCAGCGGCAA